ON THE SEMICONDUCTOR THEORY OF BARRIER-LAYER AND POINT-CONTACT RECTIFIERS

W. Schottky

Translation of a Report of the Zentral-Abteilung der Siemens & Halske A.G., Berlin-Siemensstadt, 80 pages.

(NASA-TT-F-14708) ON THE SEMICONDUCTOR THEORY OF BARRIER LAYER AND POINT CONTACT RECTIFIERS W. Schottky (NASA) Sep. 1972 81 p CSCL 20L

N73-11766

Unclas G3/26 47238

RECEIVED NASA STI FACILITY INPUT BRANCH

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546 SEPTEMBER 1972

Reproduced by
NATIONAL TECHNICAL INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

84 P

Report of the Zentral-Abteilung
der Siemens & Halske A.G.
Berlin-Siemensstadt.

ON THE SEMICONDUCTOR THEORY OF BARRIER-LAYER AND POINT-CONTACT RECTIFIERS

by

W. Schottky .

In the new theory of the unipolar processes at the boundary between metal and semiconductor, the semiconductor no longer plays the role of the poorly emitting electrode of a diode arrangement in which both thermionic electrodes are separated by a potential barrier (Section 1) but appears, like the vacuum in a thermionic tube, as carrier of a (negative or positive) electron gas of unequal density between the two metal electrodes where this gas obeys, however, other laws of fluid dynamics than in the vacuum (Section 2). At greater layer thickness, this analogy to vacuum is no longer adequate. By reason of the thermally dissociable defects existing in the semiconductor, the latter behaves similar to a space filled with incandescent sodium vapor between two thermal-emission electrodes. Particular transfer resistances at the boundary between metal and semiconductor are created when the thermal surface density of the electrons is smaller than the electron density within the neutral simiconductor determined by the volume of defects. There is then formed a spacecharge boun-

dary layer which expands and thus increases its resistance under electron motion into the semiconductor whereas it contracts to zero in the reverse direction of flow (Section 3). This interpretation of rectifier action which leads to transfer resistance and unipolar conductivity even with completely uniform distribution of defects is further expanded in Sections 4 and 5 through consideration of the additional influence of a non-uniform distribution of defects. In particular, there is discussed an impoverishment in the defects of the semiconductor at the metal boundary ("Chemical barrier layer"). It is pointed out in general that any formation of a space-charge zone within the semiconductor caused by a specific distribution of defects or through boundary effects is the cause of non-linear and unipolar conduction processes in weak fields. After once more summarizing the assumptions of the space-charge theory in Section 6, the subsequent sections attempt to give an initial review on their importance for the entire existing observational material. In all cases, the layer thickness (about 10^{-6} to 10^{-3} cm as a function of conductivity of the defect semiconductor) resulting from the space-charge theory appears to correspond to the actual dimensions of the anomalously conducting regions at the boundary between metal and semiconductor. This can be demonstrated (Section 7) for surface rectifiers from measurement of capacity and for point-contact rectifiers through an analysis of the surfaces of contact effective as "points" where generally no use is made of the assumption of a particular barrier layer at the

point of contact between point and semiconductor. Section 8 treats, from similar viewpoints, "artificial barrier layers" and also the processes in pressing foil together, under loose contact of different semiconductors with each other, and under the occurrence of interference zones ("Stoerzone") in the interior of solid semiconductors. Section 9 furnishes a review of the early history of the theory and its status within the frame of other attemps of explanation existing today.

INTRODUCTION

A preliminary communication in ref. 1 developed the basic outlines of a general theory of barrier-layer and point-contact rectifiers. The theory is an attempt to furnish a uniform explanation based on the simplest and most natural assumptions possible of the enormous multiplicity of the anomalous conductive effects in point-contact and dry rectifiers observed since the basic discoveries of Braun in 1874 (ref.2).

The present communication now discusses the principles and application of the theory in greater detail. I have here initially preferred an inductive representation supported by illustrations to a mathematical development of the theory. The circle of colleagues interested in the physical processes in point-contact and dry rectifiers and their technological appli-

cation is relatively large so that I hope to make understandable in this manner the most important basic ideas of the new concept also to such readers which lack the time and patience to follow mathematical expositions.

The exposition is based on the generally familiar knowledge of thermionic physics and investigates in the first five chapters the concepts or ideas from thermionics which so far have been transposed to the problem of electron transfer between metal and semiconductor and the extent to which these earlier concepts must be replaced by other thermionic analogies. The nucleus of the new theory is probably found in the considerations on expansion and contraction under current passage of a "boundary layer of impoverishment" occurring at the metal electrode in the semiconductor. This process dominated by space-charge conditions is investigated in great detail in Section 3 from the viewpoint of its physical nature. Section 6 briefly groups the wider and narrower suppositions of the space-charge theory and the last paragraphs are devoted to a comparison with the existing observations as well as some remarks of a historical nature.

1. The Potential-Barrier Theory. The Semiconductor as Electrode.

For interpretation of the processes in surface and point-contact rectification at the boundary between metal and semiconductor, it will be useful to utilize the experience gained in

the more easily accessible and thus more fully investigated field of thermionic physics.

A first attempt of this kind was the supposition that the metal corresponds to a freely and the semiconductor to a poorly emitting hot cathode and that the two confront each other at a distance of the thickness of a few atomic layers. This supposition made it possible to assume that, depending upon direction of flow, the free or poor emissibility of the respective cathode determines the intensity of current passage. We thus obtain rectification (ref. 3) basically determined by the work function of the two electrodes and further also through the (hypothetical) electrode gap and the local gradient of the surface forces (form of the "dish edges").

According to this theory, there is supposed to be a separation or gap between these two electrodes which is maintained by a "distance-guarding layer" formed perhaps from insulating atoms. Since the electrons in the interspace between the two electrodes must pass through a region of high potential energy, this was designated as a "potential barrier" existing between the two electrodes.

The theory of electron transfer over or through such a potential barrier (tunnel effect of wave mechanics) has been discussed and expanded in various directions. However, under

the impression of measurements of resistance and capacity made on industrial cuprous-oxide rectifiers, I felt compelled, however, already in 1929 to abandon this entire concept (ref. 4). According to the gap theory, we should have expected very high capacities corresponding to an "electrode" gap of an atomic order of magnitude but the capacities observed pointed, under any relatively non-oriented evaluation, to effective distances of 10^{-5} to 10^{-4} cm. The dependence on potential of the capacity as well as the magnitude and temperature gradient of the resistances observed made it probable moreover that the entire phenomenon concerns a process who owes its basic characteristics to the conduction and not the emission properties of the semiconductor. This supposition was strongly strengthened after the hypothesis of the defect-conduction mechanism proposed in this connnection in 1935 was confirmed experimentally in cuprous oxide (ref. 5).

2. The Electron-Gas Theory. The Semiconductor as "Vacuum"

In the meantime, the extensive investigations of Pohl, Gudden et al. on electron mobility in salt crystals as well as the extension of the Sommerfeld-Bloch theory of metal electrons to semiconductors and insulators (ref.6) had produced the finding that electrons, once they have entered into an ion crystal, travel in the latter similar as in metal. They possess a mean free path of about 10^{-6} cm at room temperature determined by

them under the influence of an electric field strength is proportional to their number, charge and, at not excessive field strength, to a constant mobility determined by the mean free path. Moreover, such free electrons behave in a crystal lattice like particles of a gas whose molecules possess approximately the mass of the free electrons and a Maxwell distribution of velocity determined by the temperature of the crystal.

These findings together with the failure of the gap theory suggested correlating the processes in metal-semiconductor systems to thermionic processes in a different manner than formerly*. It became possible to suppose that the semiconductor plays less the role of an electrode than that of a vacuum or vacuum-like intermediate in the thermionic tube. Just as a very definite electron density prevails in the vacuum above a hot cathode in thermal equilibrium, it was also possible to assume that the density of the electron gas in equilibrium in the semiconductor is linked to a work function from the adjacent metal into the semiconductor as a function of temperature. If this density of equilibrium did no longer exist at the moment of contact between metal and semiconductor, it would then occur through the fact that the metal absorbs the electrons issuing from the semiconductor and in turn emits electrons into the semiconductor.

---

* Cf. Preliminary note of author on semiconductor theory of barrier-layer, Die Naturwissensch.26,843,1938;also the study published since by N.F.Mott,The theory of crystal rectifiers,Proc.Roy Soc.London (A)171,27,1939,Further literature ref. below.

Not until the electron density is so high that the number ofemitted and absorbed particles is the same, will equilibrium be recovered. If a semiconductor layer is bordered on both sides by different metal electrodes, then different concentrations of equilibrium will prevail on the two sides. In conjunction with the fact still to be treated in detail that such concentrations of equilibrium expand or contract over the semiconductor under current passage as a function of direction, it became possible to expect non-linear and direction-dependent resistances upon current passage in such an arrangement.

The following pictures will serve to illustrate this and other thermionic analogies to be discussed further.

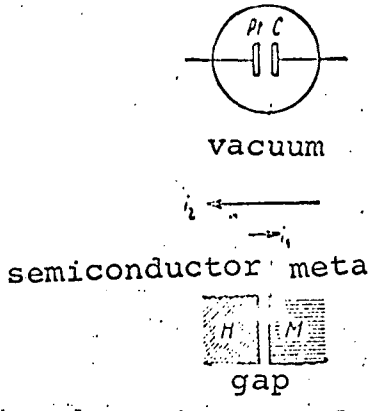


Fig. 1 - Diagram of gap theory. Semiconductor as poorly emitting electrode.

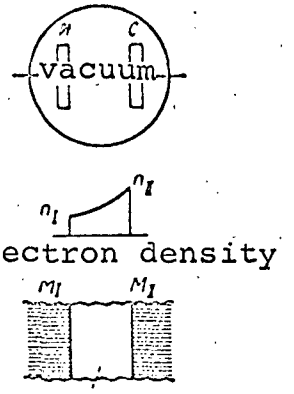


Fig. 2 - Diagram of electron-gap theory. Semiconductor as vacuum.

Fig.1 represents the earlier gap theory. The semiconductor H and the metal M here constitute the parallel to the poorly

emitting platinum electrode and the freely emitting carbon electrode of a vacuum arrangement brought to incandescent temperature. The arrows indicate the direction of current in free and in poor electron transfer. As already stated, this analogy was found to be unserviceable *.

Let us therefore turn to fig. 2 which illustrates the electron-gas theory of the semiconductor just discussed according to which rectification takes place through the different work functions between metal and semiconductor at the two electrodes. The semiconductor here plays a similar role as the vacuum between two hot cathodes Pt and C (top of fig.2). This is due to the fact that, under thermionic emission into the vacuum as well as at room-temperature equilibrium between metal and defect-free semiconductor, free electron gas exists at the density n_I over one of the electrodes and a density n_{II} over the other electrode. The distribution of density in the intermediate region is indicated by the curve in fig. 2. Just as the difference of the densities n_I and n_{II} is maintained in the thermionic vacuum by a differential of potential between the surfaces of the two electrodes designated as Volta potential, we must assume such a potential also in the defect-free semiconductor and this will be discussed in detail in future publications. However, by applying a field, we do not

---

* Cf. Historical summary, p.9

obtain, as in the true vacuum, the saturation current of the semiconductor or electrode acting respectively as cathode but a conductivity of the semiconductor layer dependent on the densities n_I and/or n_{II} i.e. the density of equilibrium of the respective metal emitting electrons is finally distributed over the entire layer as is shown when the theory is carried to its conclusion *. When neglecting effects of field emission, we arrive at limit conductivity in the forward and inverse direction of flow which have a ratio of $n_I : n_{II}$, and consequently at a rectifier efficiency as a function of this ratio.

However, even fig. 2 is not sufficient for illustrating the true relations in the known surface and point-contact rec= tifiers. The semiconductor is not only carrier but also generator of an electron gas, at least at higher layer thickness. Before we discuss this generalization of the basic conceptos (Section 3 to 5), we must return for a moment to fig. 1 and 2 for both historical and thematic reasons.

---

* This process which cannot here be followed in detail, is already very similar to the extension of a space-charge boundary layer, to be described in the next Section, under the action of an applied field. However, in the present case, the space charge does not yet necessarily play a role. In particular, the field intensity may be regarded as constant within the layer and the density distribution merely as the result of the interaction of field and diffusion currents at a given distribution of potential.

Even without the aid of the generalization of the subsequent chapters, it is possible to design, from the arrangement in fig. 1 of the semiconductor as electrode and in fig. 2 of the semiconductor as vacuum, an arrangement reproducing many features of the barrier-layer rectifiers (cuprous-oxide or selenium contact rectifiers). (In the combined arrangement of fig. 1 and 2, the semiconductor acts in part as electrode and in part as vacuum of a thermionic electrode arrangement, i.e. a defect-poor layer ("chemical barrier layer" cf.Section 4) immediately adjacent to the metal is treated as the vacuum of fig. 2 and a defect-rich region of the semiconductor immediately adjacent is treated as a quasi-metallic electrode M_{II} of fig. 2.

If we assume this "vacuum layer" as so thin and the electron density at its boundaries so minor that no appreciable space charges occur within this layer, we then arrive at the image of the barrier-layer rectifier recently proposed by Mott*. For a technical discussion of these assumptions, compare the end of this section as well as Section 5 and 7. We are here initially concerned with the fact that the author treats the defect-free layer of the semiconductor as carrier of an ideal electron gas with Maxwell distribution and well-

---

* Also elsewhere.The paper was published in the issue of May 1939 of the procedings but was sent to me in manuscript form by the author already in Feb.

defined mean free path, short in relation to the layer thickness, i.e. in analogy to case 2, but continues to speak of this part of the semiconductor as a potential barrier in the sence of fig. 1. This is a particularly conservative assumption because this makes it possible to link up woth a concepto of the potential barrier already considered by Nordheim in 1934 for the case of the cuprous-oxide rectifier in which the potential barrier between metal and semiconductor (to be penetrated wave-mechanically by the electrons) was already identified with an oxygen-poor boundary layer on the parent cooper (cf. historical review in Section 9). According to Mott the only difference from the early assumption is that the electrons do not pass through but over the potential barrier.

Clarification of these assumptions evidently requires full comprehension of the term potential barrier and the value of this term in the generalized theory. Consequently, we shall briefly also enter into the potential gradient in such an arrangement after previously considering in detail only the gradient of the electron density between M_I and M_{II} in fig. 2. In contrast to the electron density which is a completely uniform and unique physical parameter, the "potential" in question here is produced by the interaction of electric and structural influences and the determining factor is only the differential potential but not the absolute value of the potential*.

---

* In as far as this is not again uniquely established by a given physical state of reference.

The only parameter of this kind given a priori for a metal-semiconductor arrangement is the work function metal-semiconductor (or metal-insulator) which is interpreted, entirely in analogy to the work function metal-vacuum, as the difference of energy of that level of energy in the semiconductor at which the (negative or positive) electrons are located as against the level of the free energy of the electrons in the metal which we know to coincide with the "Fermi level"*

Since, on the assumption of thermal emission, the work function and the boundary density have a universal fixed relation to each other inclusive of all constants, the indication of a work function ψ_1 is completely equivalent to the indication of a boundary density n_I at one metal boundary and also on the other. An assumption of a potential barrier which would utilize no other basic elements than the work function, would consequently be completely equivalent to a theory starting from definitely given boundary densities. However, for the form of the potential

---

* A clear comprehension of this concept of the work function between metal and semiconductor became possible after the conception of the gaseous behavior of electrons and the unique definition of the energy level of electrons and semiconductors had been clarified by the work of Wilson and Peierls in 1931/32; proposals for introduction of this concept are since found in various authors. In a later publication (cf.the conclusion of this study), reference will be made to possible theoretical statements on this work function, specifically its composition by a structural and an electrostatic component corresponding to the vacuum boundary.

barrier, the work function initially gives only the height of an (almost perpendicular) steep wall at the metal boundary and consequently the level of a "high plateau" in the semiconductor adjacent to the metal. The manner in which this plateau descends or rises in relation to the interior of the semiconductor depends, like the distribution of electron density, not only on the work function (and/or boundary density) on the other side but in a generally rather complicated way*, on the magnitude and direction of the traversing current where the defect existing in the semiconductor may play a determinant role (Section 3 to 5).

It therefore seems that, when applied to a semiconductor layer with a thickness of many mean free paths, the old concept of the potential barrier loses much of its original meaning. We should perhaps retain from earlier representations of energy profiles only those which have reference to the immediate (atomic) environment of the metal-semiconductor boundary **

---

* For the no-current state, there exists, however, the familiar simple inter-relation between the electric potential and the logarithm of electron density (Boltzmann relation). Consequently the thermal distributions shown in figs. 3, 4 and 7 to 10 assuming the density to be plotted logarithmically, simultaneously represent the electric-potential profiles within the semiconductor and/or their negative or defect electrons. The potential distributions themselves are plotted in fig. 5&6.

** In addition to the concept of a plateau for this near-profile gradient initially entirely adequate for the case of purely thermal emission, we will assign, with stronger fields, to the concept of a "gable roof" profile slanted toward the interior of the semiconductor by an additional constant field, a similar significance as in the theory of field emission in vacuum.

Evidently, these considerations are not an objection to the technical part of the theory by Mott. What is concerned technically is only the question whether the anomalous drop of potential occurring in rectifiers at the effective metal electrode is created within a region firmly limited by chemical conditions or whether the dith of the anomalous layer is determined by electric effects varying with current passage. The measurement of capacity mentioned earlier as well as more recent measurements on selenium rectifiers appear to speak non-ambiguously within the meaning of the second assumption and consequently in favor of the defect and space-charge theory to be developed in the subsequent chapters.

Nevertheless, the theory is of interest as an important limit case. One of its particular advantages is moreover the possibility of a closed solution* of the differential equations involved inclusive of consideration of the effective field emission through overcoming of the image-force action** in the inverse direction.

---

* Beyond the Mott calculations, the theory can be carried through in closed form also for the forward direction if we assume the electron densities n_I and n_{II} as definitely given at the two boundaries of the ^{I} defect^{II} free layer. As noted above (by neglecting field emission; source p.372) boundary conductivities in the forward and inverse direction with a ratio of $n_I : n_{II}$. This case was calculated by the author in August 1934 in a note for internal company use "On the transition of electrons through crystal lattices without characteristic conductivity.

** A more accurate performance of the image-force calculation will require a certain modification of theorem utilized by Mott.

3. The Boundary-Layer Theory. The Semiconductor as Space with Dissociable Defects of Constant Density

We now return to the general question of the thermionic analogies to be used in the semiconductor theory of rectifier action. In relation to the case considered so far in which the semiconductor or the defect-poor layer of the latter appears only as carrier of an electron gas without inherent conductivity and can therefore be compared with the vacuum between two thermionic electrodes, we now consider the influence of the defects existing in the semiconductor and emitting electrons through the following comparison.

We continue with the mental picture of the two electrodes Pt and C in the thermionic-electrode vessel brought to high temperatures but now the interspace between them shall be filled with a ionizing gas of constant densities, e.g. sodium vapor. If we select conditions so that the electrodes do not absorb any sodium particles, their work function will remain unchanged. Consequently immediately above the two electrodes, there exist as before the electron densities n_I and n_{II} in thermal equilibrium. However, in the center between the two electrodes, the electron density is now determined not any longer by the thermal emission of the electrode metals but through the intrinsic dissociation of the soidum vapor in conjunction with the neutrality condition. We shall designate this inherent density

of the electrons dependent only on the sodium vapor pressure and the temperature as n_H. If we assume that n_H is greater than n_I and n_{II}, we then obtain, for the case of thermal equilibrium, the gradient of electron density shown in fig. 3.

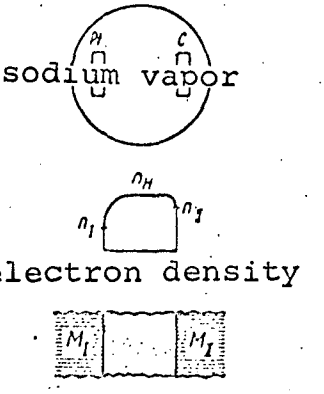


Fig. 3 - Homogeneous defect theory. Semiconductor as region with thermally dissociated defects.

Boundary layers are formed on both electrodes in which the density rises from n_I to n_H and/or drops from n_H to n_{II} whereas the constant electron density n_H prevails in the center.

We now postulate these exact same conditions in thermal equilibrium for a semiconductor with spatially constant defect density, e.g. cuprous oxide with excess oxygen or zinc oxide with excess zinc, which is located between two different metal electrodes I and II. In this arrangement, boundary layers of different electron density will be formed on the electrodes

whereas, at greater distance from the electrodes, the inherent density n_H (H = semiconductor), particular to the respective semiconductor and determined by the thermal dissociation of the defects and the neutrality condition, will prevail.

The thickness of these boundary layers is determined -- this is probably the most important property of such a peculiar physical structure -- through space-charge effects and is dependent on the magnitude and direction of the traversing current. This is explained in greater detail below. However, as a basis for understanding the phenomena in current passage, we must first consider in greater detail the no-current case, that of thermal equilibrium. In thermal equilibrium, the electron diffusion conditioned by the concentration gradient between n_H and n_I will evidently be arrested by the fact that a potential difference forms the electrons from the boundary to the interior of the semiconductor. This potential difference demands a boundary field which must have become attenuated at a sufficient distance from the edge. There is consequently required a field strength, a field-strength divergence, increasing in the direction of the boundary which presupposes the existence of a space charge. This space charge is created because the same number of positive and negative particles no longer exists at the boundary, in contrast to the interior of the semiconductor. For the thermionic boundary zone in the sodium vapor, the circumstances are easily comprehended. The number of charge particles of both

signs originating per second through thermal dissociation in a volume element is given by the (constant) vapor pressure of the neutral sodium particles. The number of particles vanishing per second is proportional to the product from the concentration of free electrons and Na^+-ions. Consequently, in thermal equilibrium, the product $n^+ x\ n^-$ is constant. The densities n^+ and n^- can individually assume entirely arbitrary values without disturbing the thermal equilibrium. Specifically, the electron density may be given by the equilibrium with a thermionic electrode as the value n_I. If this value is such that a value of n^+ greater than n^- can be calculated from the thermally determined product $n^+ x\ n^-$, then this circumstance provides for the occurrence of a space-charge density (positive in this case). Further discussion will show that the difference between n^+ and n^-, and consequently the space-charge density, decreases from the boundary toward the interior and that $n^+ = n^-$ at sufficient distance from the boundary. The electron density associated with this "quasi-neutral" state is the parameter n_H independent of electron properties which corresponds to the intrinsic dissociation of the sodium vapor and is plotted in the density curve of fig. 3.

Without following these postulates through in detail, we can convince ourselves that electron density, space charge, field divergence and diffusion and field effects mutually in-

interact so that the thickness of the space-charge boundary layer is completely determined by the given boundary density n_I and the sodium-vapor pressure (by which n_H is determined). It will be easily comprehensible that the boundary layer will be as much thicker as n_I is smaller in relation to n_H. On the other hand if we compare cases in which, through reduction of the density of the sodium vapor, n_H is reducedbut, e.g. by turning to other electrode materials, the ration $n_I:n_H$ is maintained at the previous value, the thickness of the boundary layer also increases. Thisis due to the fact that, with decreasing absolute density of the charge carrier, the absolute density of the space charges also decreases, the divergence of the field stength becomes smaller and the transition from neutral to non-neutral state takes place more slowly. If we finally change only n_H,i.e. consider the same electrode Pt in sodium vapor of different concentration as in fig. 4 where the further continuation of the

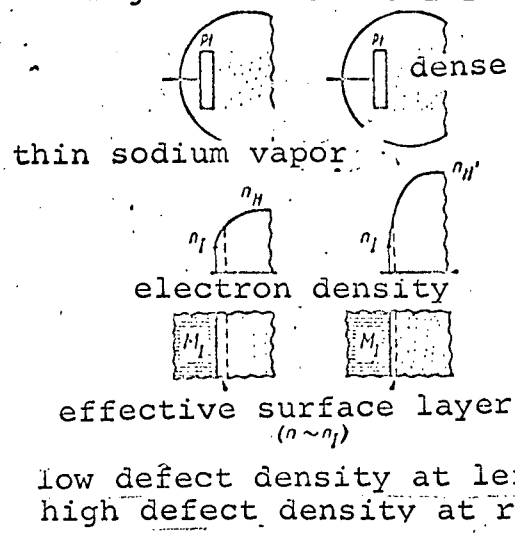


Fig. 4 - Homogeneous defects. Surface-layer thickness at different defect volume.

system is left open, the boundary layer thicknesses are shown to be of the same magnitude on the basis of the simplest assumptions in regard to defect equilibrium. However, since the rise in dense vapor within the boundary layer leads to n_H-values higher in order of magnitude, the boundary layer (n_I, n_H) in the dense vapor possesses a region narrower in magnitude with n approximately n_I than the boundary layer (n_I, n_H) in case of lesser vapor density. Under current passage, this brings about also a boundary resistance smaller in order of magnitude*.

All these considerations can now be transposed to a semiconductor with a given concentration of thermally dissociable defects. As shown in fig.4, the region n approximately equal to n_I, at the same electrode-metal (e.g. copper) contact with a defect-poor semiconductor (oxygen-poor cuprous oxide), will be much more extended than in contact with a defect-rich semiconductor (oxygen-rich curpous oxide). The considerations indicated for sodium vapor show us that the charges of opposite sign created through dissociation of the defects will play a decisive role also for semiconductors in the calculation of the thickness of the boundary layer. This offers a simplification as compared to the vapor state because we can assume the charged defect residues created through dissociation as immobile whereas, on the other hand, the possibility of the occurrence of differently charged defects may produce certain complications in comparison to the incompositely dissociating sodium vapor.

---

* The exact formulas will be furnished in a later publication

If we now send a current through a boundary layer created through electron impoverishment at the metal boundary in the semiconductor, it is initially obvious that a particular transfer resistance can be created in the boundary layer through impoverishment of conducting particles and that this resistance will even be high, at a sufficiently low ration n_I/n_H, as against the resistance of a layer of the same thickness with the normal conductivity n_H. However, only at the very lowest current loads will it be possible to expect an unchange quasi-thermal distribution of electron density in the boundary layer. If somewhat higher currents pass through, all conditions previously determining the thickness of the boundary layer change and, whereas the thermally determined* electron density n_I and far within the interior of the semiconductor the intrinsic density n_H continues to prevail, the thickness of the boundary layer -- and consequently its resistance -- changes to a very large extent through the passage of the current.

In contrast to the case of thermal equilibrium considred only superficially above, we must now examine the collaboration of the forces determining electron distribution under current density in greater detail because this furnishes the actual key for the comprehension of the unipolar processes in space-charge layers.

* Influences on n_I by fields occurring on metal electrode play large role in the further development of the theory but shall hwere provisionally not be taken into account.

We first consider the case in which the electrons are forced, by the applied field, from the electrode into the semiconductor. In order to have the simplest possible conditions we assume the applied field strength as such that the diffusion current can be neglected against the field currents. We then know that, since no accumulation of charges can be permitted to occur, the current density in the stationary state must be spatially constant (and/or free of divergence), i.e. the product of field strength E and conductivity $\varkappa$ and/or electron density n must be constant for plane boundary layers:

$E \times n$ = constant (steady-state condition for plane boundary layers).

Where low electron density n exists, there must consequently prevail a high field strength and, at points of high electron density, field strength must be low. Steady current passage through a boundary layer with variable n therefore demands a spatial change of field strength in the boundary layer which is inversely proportional to the change of electron density.

For example, if the n-distribution in the boundary layer is given by the gradient designated in fig. 5-a then the potential gradient must have about the form given in fig. 5-b. The potential must increase from the electrode I (which is here connected as cathode) to the semiconductor interior (which acts as anode of the partial process) in such manner that the

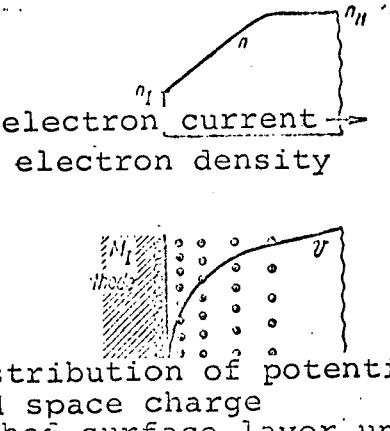


Fig. 5 - Impoverished surface layer under electron transfer from electrode to semiconductor.

potential increase and consequently the field strength has the gratest value(amount) at the boundary I where the smallest density n_I prevails, and possesses, in the semiconductor interior with the highest density n_H, its lowest value which continues to remain spatially constant.

The deciding factor now is that, to the potential gradient shown in fig. 5-b, there must necessarily be linked a very definite space charge by reason of the Poisson equation. We know that a positive space-charge density corresponds to the potential curve convex downwards in fig. 5-b. This can also be

expressed so that, of the many lines of force terminating in the negative surface charge of the metal cathode, only a small part comes from the semiconductor interior but an appreciably greater part issues from the positive space charges of the boundary layer and terminates in the cathode. Fig. 5-b indicates the distribution of these positive charges in accordance with the potential curvature shown.

However, in order to make these different requirements for the boundary layer compatible with each other, it is necessary for the boundary layer which is primarily a region of electron impoverishment, to become also a region of positive space-charge density. This is actually accomplished by the interaction (unique in quasi-thermal electron motion) between electron density and resulting space-charge density. A positive space-charge density in the boundary layer is determined both by the absence of negative (electron) charges and also by the increase of the positive (Na^+) charges as against the neutral state as demonstrated by the example of sodium vapor.

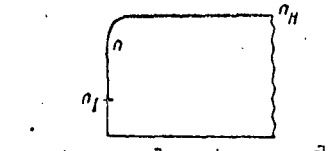
We therefore see that the distributions of particle density, potential, and space charge shown in fig. 5-a and 5-b are compatible with each other and with a steady-state condition. Even without diffusion effects, a steady distribution of the electron densities and the space charges can develop. We see without difficulty that, in this region of "pure field

emission" where current density, with given n, is proportional to field strength, the field strengths themselves and consequently also the differences of field strength between boundary and semiconductor interior must increase proportional to current density. This means, however, that the total number of the positive space charges distributed per square centimeter in the surface layer must also increase proportionally with the current. Since the density of the space charges is definitely determined by the electron density confined in the fixed limit n_I and n_{II}, the additional space charges can be obtained only through enlargement of the space-charge zone and consequently of the impoverishment layer. Let us summarize this:

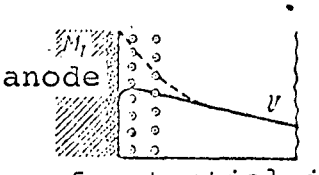
If, in an impoverishment layer, the electrons flow from the impoverishment electrode into the interior of the semiconductor, there is then constituted a steady electron distribution in the surface layer in which the space charges produced by electron impoverishment themselves provide for the development of the differences in field strength (required by reason of the different electron densities). The width of the surface layer continues to increase with increasing current density and it is evident that the causes of this increase are effective already at low current intensities so that we must therefore assume, starting with the diffusion layer constituted in the state of equilibrium, a steady increse of the surface-

layer thickness and consequently also of the resistance of the surface layer in the inverse direction.

Circumstances are quite different when the electrons flow from the interior of the semiconductor to the electrode (electrode I as anode as in fig. 6-a and 6-b). If there existed here



a electron density



distribution of potential in space charge

Fig. 6 - Impoverished surface layer under electron transfer from semiconductor to electrode.

also an extended surface layer similar to fig. 5-a, there would then have to prevail in this surface layer, in order to maintain the product E x n constant, an appropriate potential distribution such as is represented in fig. 6-b by the broken-line curve, i.e. such a distribution could be produced only through (extended) negative space charges (potential curvature upward). However, since there exists, on our premises, a

priori electron impoverishment, i.e. a positive space-charge density in the surface layer, the broken-line potential gradient of fig. 5-b is not possible. If a curvature of the potential curve takes place in the surface layer, it can be oriented only downward as represented in the solid-line curve. Accordingly there must exists, in the interior of the semiconductor, lower field intensities instead of lower electron densities and possibly even opposing fields. A steady state of flow with pure field flow for the surface layer is therefore not possible in this current direction. What actually occurs is instead that the electron density n_H of the semiconductor extends unchanged as far as the electrode boundary and that the surface layer with n not equal to n_H contracts so (fig. 6-a) that the diffusion curves--correspondingly increased through the enormous increase of the concentration gradient -- now are able to take over the task of carrying the electrons, with insufficient and even oppositely oriented field effect, through the surface layer to the electrode. Similar considerations which enabled us to understand the increase of surface-layer thickness proportional to current density in the electron direction from metal to semiconductor, indicate for the electron direction from semiconductor to metal that, at large current densities, the surface layer contracts inversely proportional to current density.

This rounds off the overall picture. The contraction of

sign in the impoverishment layer and consequently the forward direction is again the one in which the conducting particles flow toward the electrode. In all cases where the electrode creates a surface impoverishment of the type of particle funishing conduction in the semiconductor -- and only these cases lead to measurable transfer resistances -- there is then valid for the direction of orientation of the rectifier action of a surface layer the following easily memorized rule: the mobile particles prefer to flow from the semiconductor into the metal rather than inversely. In this formulation, the action conditioning the barrier resistance is most clearly linked with its cause, i.e. that the metal "replenishes" the mobile particles for the semiconductor less well than the semiconductor itself.

The postulate here developed for the surface layers at the boundary between metal and semiconductor (with homogeneous defect density) was communicated first in 1935 as empirical rule for the direction of rectification in excess and defect semiconductors*.

The fact that it has since been confirmed in further

---

* W. Schottky, Discussion to the lecture by Störmer, ZS,f.techn. Phys.16,512,1935. One half of this rule (direction of rectification in defect semiconductors) is already contained in the theory of the cuprous-oxide rectifier proposed 1929 and has played a large role in the correspondence between Peierls and the author in subsequent years.Cf. page 408 (of source).

small and no oppositely acting surface layer occurs on the other electrode), the boundary conductivities in forward and inverse direction will obviously have a ratio of n_H; n_I so that the "efficiency" of rectifier action is determined directly by this ratio or its reciprocal value, the relative impoverishment. (Field effects in the surface-layer electrode are here not yet taken into account).

In addition to this concept of the expansion and contraction of a impoverishment surface layer, we can illustrate this as well as the current-dependent influences on density of the electrons in semiconductors to be discussed later, by the concept of "electron drift". We can state in general that the conductivity existing in the case of no-current always "drifts" through current transport in the direction in which the conducting particles flow. Consequently, the "natural" surface-electron impoverishment drifts into the semiconductor in the inverse direction but, in the forward direction, the surface layer is finally completely "clogged" by the conduction electrons of the semiconductor.

If the current in the semiconductor is not carried by ordinary electrons but by defect electrons, i.e.by particles hehaving like positive particles with positive mass similar to the electron mass, then the forward and the inverse direction is reversed because of the change of the space-charge

the surface layer (which we must assume based on high current densities in the direction of motion from impoverishment electrode to semiconductor) does not stop at zero potential where the surface layer has its natural thickness conditioned by equilibrium of diffusion, but continues without limit, under increase of the current in the inverse direction of motion, from semiconductor to impoverishment electrode. Accordingly, the resistance of the surface layer has its highest value for the direction of electron flow from electrode to semiconductor, becomes increasingly less with decreasing potential, and practically vanishes in the direction of electron flow from semiconductor to electrode at higher current densities.

An electron-impoverishment surface layer possesses a unipolar conductivity whose inverse direction is given by the electron transfer from the electrode into the semiconductor whereas the barrier resistance vanishes in the direction of electron flow from the semiconductor into the electrode (forward direction).

How high the differences of conductivity in forward and inverse direction made possible by this mechanism may be, can be estimated even without detailed discussion. Under the most favorable conditions (i.e. if the "series resistance" of the semiconductor not influenced by the surface layer is sufficiently

verifiable cases * indicates that a similar process as the surface impoverishment described here, in the electrode controlling rectification decisively participates in actual rectifiers. A direct proof for changes of thickness of an effective resistance layer -- expansion in the inverse direction -- can also be found in the already mentioned trend of the rectifier capacities with the bias voltage. It will therefore be accepted as justifiable if I subject in future publications the behavior of an impoverishment surface layer at the boundary of a semiconductor with homogeneous defect content to detailed examination and also compare the more specific findings to be derived from this result with the observations in actual rectifier arrangements.

In order to be able to evaluate at least to some extent the significance of such comparisons, we must continue preceding qualitative considerations somewhat further. We will want to know first of all the role of the second electrode. Circumstances are here very simple if the second electrode is sufficiently large in relation to the first as is the case in most point-contact detector arrangements. With homogeneous defect content of the semiconductor and the same electrode material, we then have two oppositely connected qualitatively equal surface layers where, however, only the surface layer formed at the locus of point contact in which the current lines concentrate, is

---

* For zinc oxide by W.Hartmann, ZS.f.techn.Phys.17,436,1936; for titanium dioxide by H.Klarmann, Wiss, Veröffentl.Siemens-Werke 18,2.H.,198,1939.

controlling for the resistance (diffusion resistance). Another simple ideal case exists if the second electrode is only as large as the first but is assumed as impoverishment-free ($n_{II}=n_H$ or, in any case, n_{II} very much greater than n_I).

4. The Semiconductor as Space with Non-Uniform Defect Density. Definition of "Barrier Layer"

Of greater interest than these cases of the "non-blocking" second electrode for reasons of geometry or emission is a rectifier arrangement with two equal and equally large electrodes (e.g.copper) in which the surface-layer effect on both electrodes is of different order of magnitude because the defect density (oxygen content of cuprous oxide) -- assumed as homogeneous within each surface layer -- and consequently the intrinsic conductivity n_H in the vicinity of both electrodes is of different orders of magnitude. We would then have the two partial arrangements of fig. 4 so-to-speak oppositely connected.The question then is only whether the (steady or non-steady) transition from low to high defect density within the semiconductor may not lead to new rectifier effects which may possibly overlap the surface effects.

The question thus posed of the current-voltage relations in semiconductors with non-uniform defect density compels a

further generalization of the analogies represented in fig. 2 to 4. Here it will again be appropriate to initially regard the case of thermal equilibrium without applied potential. The top of fig. 7 represents a vacuum vessel (heated to high

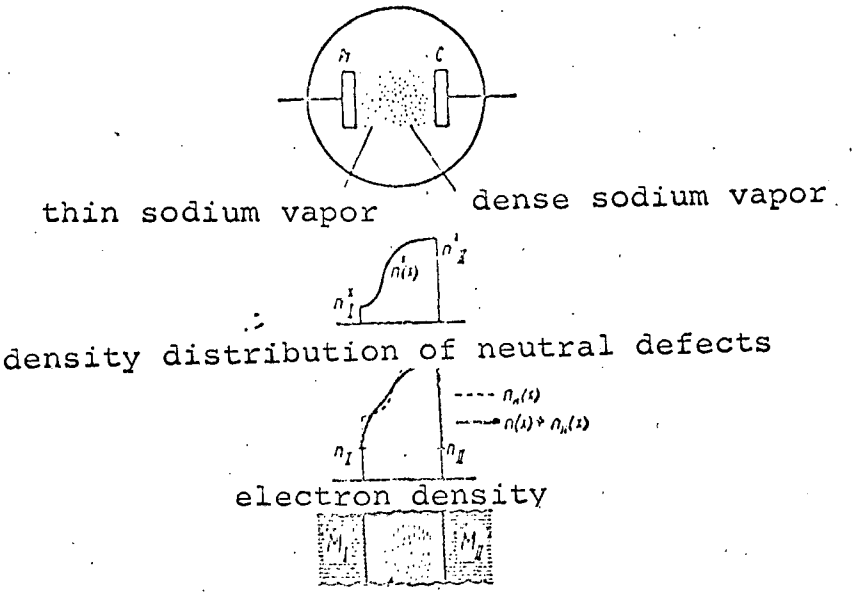


Fig. 7 - Non-uniform defect distribution, surface and internal effects.

temperatures) with two arbitrary electrodes,e.g. Pt and C, again with sodium vapor between them but now in non-uniform distribution. This non-uniform distribution of density n^x of the sodium vapor (x = neutral sign) maintained constant by some artifice between the electrodes, is represented by the

curve path below the thermionic electrode diagram. In the curve below this diagram of the distribution of electron density under thermal equilibrium, the broken-line curve indicates the density which would result if the quasi-neutral intrinsic density n_H of the electrons associated with the vapor density n^x existed verywhere. The solidline curve of the actual density distribution shows both the already familiar surface deviations due to the surface effects n^I and n^{II} given by the electrode emission and also deviations in the interior region. (Let us assume all densities to be plotted in a condensed logarithmic scale so that even small deviations have a high significante). These are deviations in the sense that electrons overflow from a region of high intrinsic density to a region of lower intrinsic density so that the true electron density n is less than n_H on the defect-rich side of a concentration gradient of the density n^x and is greater than n_H on the defect-poor side of the gradient. Ahead of the electrode, there is then again traversed, under the effect of surface impoverishment, the neutral point $n = n_H$ and the n-curve drops to n_I.

These relations are now transposed unchanged to the semiconductor region (botton of fig. 7). Between two arbitrary electrodes M_I to M_{II} in the manner described above. The (defect) electron density produced in equilibrium will then also have the gradient represented, i.e. deviations from the quasi-neutral

state as a consequence of non-uniform defect density can be understood for the case of thermal equilibrium as entirely similar to the surface layer. The difference due to non-uniform defect density of n_H in adjacent zones would then require a corresponding potential difference for its maintenance. The potential "profile" thus associated with the given n^x-distribution would have as prerequisite a given space-charge distribution in the interior which can be accomplished only through deviations between n and n_H. The actually resulting density n represents a sort of compromise between the attempted intrinsic density $n_H(x)$ and the possibilities, limited by the space-charge densities, of maintaining this density at the locus x in electrical equilibrium.

When applying an additional exgernal field, electron density will change in all not quasi-neutral regions of the semiconductor. Poorly conducting regions lying beyond freely conducting regions in the direction of motion of the conducting lying b particles are "clogged" with conducting particles, i.e. are increased in their conductivity; freely conducting regions lying beyond poorly conducting regions will be impoverished of mobile charge carriers. Any occurrence of non-quasi-neutral regions, i.e. of space-charge zones within a semiconductor of non-uniform defect distribution, must therefore be regarded as the cause of nonlinear resistance and specifically as cause of unipolar conductivity. Inversely, regions

in which the quasi-neutral state is practically still conserved even under current passage will act as (locus-dependent but) pure ohmic resistances. This general rule furnishes a perhaps obvious but in any event very important criterion for evaluation of all non-linear processes which may occur with weak fields in semiconductors. With strong fields (about 10^5 Volt/cm and beyond), new effects may occur. Specifically, the electron velocity produced by the field may be of influence on the current-voltage relations so that even quasi-neutral regions may show non-linear resistances. However, the clogging effects occurring in non-quasi-neutral regions with the weaker fields will always have to be regarded and considered a priori as the most important cause of non-linearity.

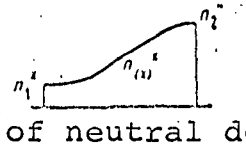
Quantitative pursuit of these considerations shows that, depending on the defect content and the intrinsic conductivity of the respective semicon ductor region, the space-charge compensation without applied field takes place in semiconductors of an intrinsic conductivity of 10^6 to 1 Ohm^{-1}/cm^{-1} within distances 1_o of an order of magnitude of about 10^{-6} to 10^{-3} cm. Under current passage, the thickness of these space-charge zones may expand or contract but their order of magnitude is fixed by these data. It follows from this that for the manifestation of internal space-charge effect, appreciable changes of the defect content within very thin semicon-

ductor layers are a prerequisite. This is a circumstance which makes the experimental investigation of these layers and the determination of defect distribution in them extraordinarily dif ficult. On the whole, we may assume that incomposite semiconductor layers, even if they have inhomogeneous defect distribution will have to be regarded for the predominant part of their volume as ohmic resistances. The non-linear effects are restricted as a rule to very small and especially defect-free zones which will lie, on the basis of creation of the defect distribution, generally at the surface and only in certain cases (carborundum) which the semiconductor investigated. To this theoretical expectancy of the space-charge theory of the semiconductors corresponds a large observational material on cuprous oxide and various detector substances which have found its expression in the postulate of thin defectfree "barrier layers" at the effective electrode. The discussion of these special cases (fig. 8 and 9) will show that the unipolarity of such a barrier layer in limit cases is based either on a pure surface-layer effect dependent on electrode emission n_I in accordance with fig. 4 or on an "internal-layer" effect independent of the electrode material. The term "barrier layer" thus receives a somewhat opalescent and vague character. I should like to propose its use, in the space-charge theory of the semiconductors generally, in the sense of "chemical barrier layer", i.e. specifically for the characterization of a defect-free (surface or internal) zone. "Barrier-layer rectifiers" would then be those in which a defect-impoverishment zone

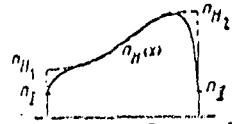
occurring in the vicinity of one electrode plays a decisive role, or, more generally, rectifiers with non-uniform defect distribution of the semiconductor corresponding to a non-uniform sodium-vapor pressure in the thermionic substitute diagram. It is here assumed further that the barrier-layer zone, disregarding its defects, has the same large structure and the same composition as the adjacent parts of the semiconductor. Layers different in composition and lattice structure from the semiconductor should preferably be designated as "heterogeneous or impurity layers".

5. Special Cases of Non-Uniform Defect Distribution - Quasi-homogeneous Distribution, Abrupt Rise, Linear Distribution.

Fig. 8, 9 and 10 represent some of the most important special cases of defect distribution which may be controlling for the formation of the chemical barrier layer and its electrical effect. The thermionic substitute diagram corresponds in every case qualitatively to the top of fig. 7.



density of neutral defects



electron density

slowly variable oxygen content

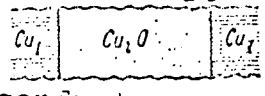
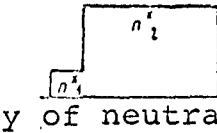
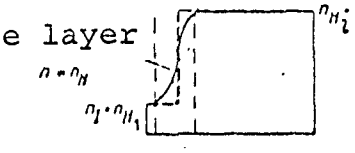


Fig. 8 - Semiconductor arrangement with quasi-homogeneous defect distribution.



density of neutral defects

space-charge layer



intermittently variable oxygen content

Fig. 9 - Semiconductor without surface layers and only internal effect.

The defect distribution in fig. 8, top curve, refers to the case considered at the beginning of the preceding section in which the defect density within the electrode surface layers can still be regarded as homogeneous. As shown by the curve of distribution of electron density plotted below this, the quasi-neutral state is assumed for the entire internal region in which neutral density n^x increases from n_1^x to n_2^x . The the extent in which this condition is satisfied even under current flow, the entire interior part of the semiconductor acts like an ohmic resistance where conductivity may be assumed moreover as so rapidly rising to the right that the entire resistance in the region n^x greater than n_1^x does not play any appreciable role in relation to the resistance of the n_1^x-zone controlled by surface effects. Because of the high intrinsic density n_{H2}, the resistance of the second surface zone can also be neglected even then when the electrode consists of the same material, i. e. $n_{II} = n_I$. The associated semiconductor arrangement is represented by the example of the cuprous-oxide rectifier with two copper electrodes described at the beginning of section 4.With this layered structure, the rectification of a cuprous-oxide layer which has the low oxygen content inherent in the equilibrium with the copper phase (ref. 7) in a certain zone on the parent-copper side and then changes to oxygen contents of atmospheric equilibrium toward the outside, could be explained as a pure surface rectifier of the system parent-copper and high-ohmic cuprous oxide within the frame of the homogeneous

surface-layer theory. Since the entire arrangement behaves electrically both on the surface and in the interior of the semiconductor so as if the defect density were locally homogeneous in sections, we may therefore speak in such types of quasi-homogeneous defect distribution.

To a certain extent, the counterpart to this quasi-homogeneous arrangement is the rectifier type outlined in fig. 9. We have here assumed a defect distribution n^x which possesses the constant value n_1^x ahead of the first and the higher value n_2^x ahead of the second electrode and which abruptly rises from one to the other value in the interval between them*. It is further assumed that the electrodes are selected so that their n_I and n_{II} by chance agree with the quasi-neutral densities n_h1 and/or n_h2. There will then exist no surface layers at all and there will be formed instead a space-charge layer with unipolar resistance properties at the locus of jump n_1^x/n_2^x. The space-charge zone in the vicinity of this locus of jump will possess a certain extent both in the freely as well as in the poorly conducting part of the semiconductor. However, as in the two space-charge layers represented in fig. 4 and created by surface impoverishment, the extent of the space-charge zone into the defect-poor part n_1^x will be greater than in the defect-rich n_2^x. The total resistance will be less if the conducting particles "drift" from the freely conducting section

* Unfortunately, this distribution is not represented entirely accurately in the lower part of the fig.

n_2^x into the poorly conducting part because then the poorly conducting section becomes "clogged" with conducting particles. On the other hand, if the particles are forced from the poorly conducting layer to the freely conducting layer, there then prevails the low intrinsic conductivity n_H1 corresponding to the density n_1^x as far as the geometric defect boundary. "In space-charge layers forming at points of high concentration gradient of the defects, the conducting particles prever to flow from the freely conducting to the poorly conducting part of the semiconductor rather than inversely"*.

The wholly internal rectifier is, however, an ideal case like the quasi-homogeneous surface-layer rectifier whose suppositions cannot generally be satisfied strictly. As a rule, surface and internal effects will overlap with higher inverse currents where the space-charge layers expand increasingly. In further development of the theory, we will therefore not be able to avoid the consideration of problems in which the defect density changes appreciably already within a space-charge surface layer and where, on the other hand, possible internal effects are modified by surface effects. It will therefore be useful to bring out also a relatively simple and characteristic type for

---

* Let us note in passing that, even at a completely homogeneous defect content, an electric double layer incorporated for some reason in a semiconductor may produce a finite space-charge disturbance and a consequent non-linear, unipolar internal-layer defect.

the general case of overall non-uniform defect distribution represented in fig. 7. Fig. 10 assumes that the defect content (e.g. oxygen excess in cuprous oxide) increases from the parent-copper electrode Cu_I to the externally mounted electrode (e.g. Cu_{II}) precisely in such manner that the intrinsic density n_H is a linear function of locus *; the intrinsic densities of the semiconductor at both electrodes are again designated by n_{H1} and n_{H2}. The density of the neutral particles (top curve) here does not show a linear trend with x but, under the simplest assumptions on the internal defect equilibria, a quadratic trend. The case considered must be regarded, however, as particularly simple and important for two reasons. On the one hand, there occurs, in the general differential equations for the distribution of electron density n in the case of inhomogeneous distribution, a local intrinsic density $n_H(x)$ directly as representative of the defect distribution. The assumption of a linear trend of n_H with x is therefore mathematically the simplest. On the otherhand, it can also be shown that a layer whose defect distribution is due to a process in which a divergence-free diffusion stream flows from the e.g. oxygen-rich to the metal rich side, has a parabolic distribution of n^x and therefore a linear distribution (even at low temperatures) of n_H. This is

* For such reason, a linear scale for the n-ordinates is to be assumed in fig. 10. However, we must imagine that the ratio $n_{H2}:n_{H1}$ may amount to many powers of tehn.

valid in any event if diffusion takes place only through motion of the dissociated particles as has been confirmed for many oxides and similar ion lattices (Ref. 8) and if, at the diffusion temperature, the formation of space-charge zones of finite extent is not yet involved. A further discussion of these conditions is, however, not possible within the frame of this communication*.

As shown by the solid-line curve of the center part of fig. 10, the quasi-neutral state has been assumed for the greater part of the layer so that the n-curve coincides with the n_H-curve. Deviations of space charge occur only on the surfaces and again predominantly on the defect-poor side. The drawing represents a special case in which n_I is assumed as so small in relation to n_{II} that the intersection of the n and n_H-curves demonstrated in the discussion of fig. 7 no longer occurs.

The behavior under current flow of the surface zones so far presented only for the no-current case does not easily become clear in its details under defect distribution as assumed in fig 10. However, certain general considerations can be made on whether the defect distribution within the space-charge zone can still be regarded as quasi-homogeneous or not. Since the space-charge zone contracts arbitrarily in the forward direction but under homogeneous defect density would arbitrarily expand in the inverse direction, we shall have to expect in general that

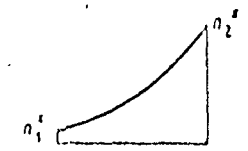
* (See next page)

the surface layer behaves under forward current and perhaps also at zero potential approximately quasi-homogeneous but finally expands under higher inverse currents in regions of highly increased defect concentration and finds in these freely conducting regions a natural and rather sharp delimination. This is precisely the behavior on which we have to conclude from the trend of the alternating-current capacity of the cuprous-oxide back-plate rectifier** .We obtain in the forward direction an apparently arbitrarily continued rise of the alternating-current capacity but, in the inverse direction, only a slight drop and, at moderate barrier potentials, practically constant values. We are therefore probably able to say that the present concept of the cuprous-oxide barrier layer which was probably more similar to fig. 9, is not necessary consequence of the capacity measurements; even under steady variability of the defect density, similar results may be expected.

---

* Cases of not quasi-neutral diffusion are conceivable where, in the steady state, the density n^+ of the ionized defects possesses a minimum at some distance from the defect-poor semiconductor boundary. At lower temperatures, this may also condition a local minimum of n^x.

**Cf.W.Schottky u.W.Deutschmann,Phys.ZS.30,839,1929.Let us point out here that it is possible, especially in the forward direction, to determine the capacity of the "barrier layer" accurately only then when, in the substitute diagram, a compensating resistor is introduced in addition to the series connection of resistance and capacitance.However,if we utilize only the parallel capacitance diagram, there are then measured, in the forward direction,very rapidly decreasing capacities becuase the interposed ohmic layer resistance soon terminates the barrier-layer resistance as much more as the frequency (measuring) is higher.



density of neutral defects

electron density
diffusion distribution of oxygen content

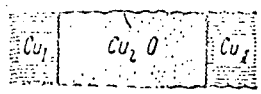


Fig. 10 - Semiconductor arrangement with linearly increasing characteristic conductivity of semiconductor

6. Summary of Assumptions

When we realise the enormous multiplicity and complexity of the process dependent on potential and direction in the field of surface rectifiers, of point-contact detectors, and of current flow through thin crystal layers as well as the large number of the different attempts at interpretation published in this field alone during the last decade, then the selection of basic phenomena made above which is now to be held responsible for the totality of these processes, will appear as rather narrow and may raise the supposition that only a new unreliable generalization is taking the place of the earlier ones.

Consequently, instead of now turning to the mathematical formulation of the theory as a follow-up to the preceding discussion, I should like to conclude this first publication on the new theory through further considerations of the general physical character which are intended specifically to afford a priori a certain insight into the experimental significance of the theory.

Let us first briefly summarize the assumptions of the new theory. This appears useful because rather different prerequisites may be concerned in accordance with the considerations in Section 1 to 5. We differentiate here between general assumptions of the theory and special prerequisites, i.e. those which we regard as particularly important for the interpretation of the actual phenomena within the frame of the general theory.

General Assumptions

1 - The semiconductors to be considered are pure electron conductors, i.e. either with excess or with defect conduction. The excess or defect electrons -- both are designated in the following as electrons for the sake of simplicity -- behave in the semiconductor with adequate approximation like an ideal gas with constant particle mass similar to the free electron mass.

2 - The ions or other charged defects existing in the semiconductor are assumed as completely immobile; ion mobility is considered only in questions which involve the antecedents or the long-time changes (ageing) of the respective semiconductors.

3 - At any point of the semiconductor, electron flow under the effect of applied fields is assumed still as quasi-thermal; the energy gain in the field within a mean free path shall be small in relation to the intrinsic thermal energy. It shall therefore be permissible to assume proportionality between field intensity and electron current, i.e.a "mobility" of the electrons independent of field strength.

4 - The semiconductor is assumed expressly as defect semiconductor*. Because of the quasi-thermal behavior of the electrons (assumption 3), thermodynamic equilibrium prevails everywhere between fixed defects and mobile electrons. The case of vanishing defect density (electron density determined only by surface effects, Section 2, "Semiconductor as Vacuum" shall be included in the theory as limit case.

---

* We shall discuss in later communications why it is not possible to construct rectifiers from "characteristic semiconductors".

5 - In the no-current state, a zone shall be created through external conditions within or on the surface of the semiconductor in which the charges of the fixed defects and of the electrons do not cancel each other. There then occurs, according to the general theory, under current flow in this zone, a resistance dependent on potential and as a rule also on direction. Since the study of such space-charge zones* in semiconductors is the subject of the general theory, it will be possible to designate this theory generally as (quasi-thermal) space-charge theory of crystal rectification.

Special Prerequisites

6 - In the arrangements investigated by the special theory, there is assumed an impoverishment of that type of electrons on at least one of the metal electrodes in contact with the semiconductor which provides the transport of electricity through the semiconductor.

7 -a - In the simplest case, there shall be given, at the metal boundary of the semiconductor, a density of thermal equilibrium of the conducting type of electron determined by the work function metal-semiconductor and independent of defects and any occurring fields. This surface density shall

---

* The expression "space charge" is not necessarily intended to characterize an appreciable influence on potential distribution but an appreciable deviation from quasi-neutral state.

be less in defect-free semiconductors than the density of equilibrium in the quasi-neutral parts of the semiconductor.

7-b - For better adaptation to actual conditions, a change of work function through the surface field shall be taken into account (Schottky emission, tunnel emission). This influence can be produced either through external fields or may exist already in thermal equilibrium as consequence of a zero field. Further local influences (point-contact defects, local effect of interposed defects, etc.) must be taken into account.

8 - The treatment of intermediate impurity layers between metal and semiconductor is provisionally eliminated from the field of the special theory. "Barrier layers" are understood to be only surface regions of the semiconductor characterized by lesser defect volume.

9 - Treated in preference by the special theory are semiconductors with (relatively low and) completely uniform defect volume as far as the surface. The electron-poor layer to be assumed at the metal boundary of such semiconductors according to 6 and 7 is designated as (homogeneous) surface layer.

In pursuit of the theory, the density of the fixed

and mobile charges is assumed as so low that corpuscular interaction of the charges need not be considered. We can therefore expect a mean space-charge density in each volume element.

In the following, interpretations utilizing only the general assumptions 1 to 5 are assigned to the general "space charge theory of crystal rectifiers". Those considerations based partially or completely on the special assumptions 6 to 9 will be designated as "surface-layer theory". Consequently, unless otherwise stated, "surface layer" means "impoverished surface layer" (assumption 6) and also "homogeneous defect volume" (assumption 9) in the vicinity of the metal electrode.

7 - Barrier-layer, Thin-film, and Point-contact Rectifiers

In the following comparison with the existing experimental material, there is obviously not yet involved a systematic investigation and even less a proof of the theoretical assumptions on the basis of the existing observational material. A valid test of the theory representing simultaneously a direct determination of the controlling constants such as work function, defect distribution, etc. does not even seem possible to me if we only experiment with so-to-speak "complete" rectifiers. It will be necessary for this purpose instead to

make an attempt at separating and measuring the individual elementary effects under the most simple conditions which need perhaps not even be related to rectification.

If we enter, under these reservations, into discussion of some important general facts observed in the field of crystal rectification, we may then a priori point to a property which is common to barrier-layer (surface) and point-contact rectifiers and is necessarily peculiar to the processes in thin films, i.e. the order of magnitude of layer thickness in which the entire non-linear resistance appears concentrated in all cases. From the already mentioned measurements of capacity in cuprous-oxi-e rectifiers and also from corresponding measurements on selenium rectifiers, it was shown that the effective layer thickness here has ian general an order of magnitude of 10^{-5} to 10^{-4} cm. As already stated, the space-charge theory leads to the same order of magnitude between 10^{-6} to 10^{-3} cm as a function of the $\varkappa_h$-value of the semiconductor adjacent to the metal. If we adopt the viewpoint that, in barrier-layer rectifiers, the effective layer thickness is determined in actuality through the thickness of the chemical barrier layer, specifically through defect impoverishment at the metal boundary dependent on diffusion processes in manufacture, this concordance of order of magnitude will perhaps be regarded as a random factor. In refutation, I should like to point here to a consideration which has played a role in the development

of the space-charge theory. There is certainly no objection to assuming some arbitrary defect distribution in the semiconductor strongly decreasing in the direction of the metal electrode. Nor can it be objectionable if we construe, by assigning to each point the associated (quasi-neutral) electron density of equilibrium, the "potential barriers" necessary for the conservation of this maintenance as well as the "strong fields" under certain circumstances linked to the potential barriers so that several possibilities remain open for the interpretation of rectifier action. However, if we go below a certain limit dependent on the defect density itself in the degree of the assumed defect change, we then come into conflict with the space-charge theory whether we regard the latter a priori for important or secondary. In the no-current state, it is impossible to produce arbitrary high and steep potential barriers within a defect (impurity) semiconductor precisely because the Poisson equation exists between potential curvature and space-charge density and the latter is limited so-to-speak "ruthlessly" by the defect content of the semiconductor*.

---

* This is perhaps best seen from fig. 9 for the particle density n which is linked to the potential by the Bolzmann relation. If the n_1^x-zone is selected small as against the thickness of the space-charge layer, then this entire zone is flooded by electrons (however in a manner still depending on n_I and n_2^x) and we no longer note the abrupt jump of n^x in the n-distribution and consequently of the potential distribution.

We can therefore state that the drop of defect density in a chemical barrier layer can be utilized for rectifier action only in so far as permitted by the space-charge conditions. However, then the dimensions following from the space-charge theory appear in fact as controlling for the lowest possible unipolar effective layer thickness also in barrier-layer rectifiers.

With equal compulsion, the space-charge theory also determines the upper limit of layer thickness in semiconductors within which anomalous conduction processes are still possible. Regions which can no longer display, according to the space-charge theory, any significant deviations from the quasi-neutral state, must have, with weak fields, a purely ohmic resistance because their electron density is determined by the quasi-neutral state condition independently of the direction of current.

If we were therefore forced to decide whether we shall regard, specifically in the cuprous-oxide rectifier, the space-charge distribution or the defect-poor diffusion layer (depending on the antecedents of the rec tifier) as controlling for the thickness of the anomalous conducting zone, we should necessarily opt in favor of the space charge. However, what is concerned is not a strict either-or. We may suppose on the contrary that the optimum of technical rectifier action is obtained under conditions where the characterizing dimensions

of the defect distribution in the diffusion layer are of the same order of magnitude as the associated characteristic length of the space-charge theory. Accordingly, the fact, initially confusing any interpretation, of the same order of magnitude of chemical barrier-layer thickness and physical space-charge thickness would be so-to-speak the result of a "technical genetic selection".

Since these considerations were historically the principal inducement for the development of the space-charge theory, they cannot yet be listed, strictly speaking, among the experimental confirmations of this theory. Let us therefore occupy ourselves now with applications of the space-charge theory to such cases of crystal rectification where the existence of chemical barrier layers is to a large extent improbable or excluded, i.e. point-contact rectifiers and rectification of thin films between metal electrodes in areal contact.

It will certainly reinforce confidence in the space-charge theory if the latter is able to predict not only the existence of these two additional rectifier types but to indicate in order of magnitude the effective dimensions of the point-contact area and/or the thickness of the thin films without making use of any ad-hoc assumptions on inhomogeneities in the respective semiconductors.

In such arrangements, the impoverished surface layer does not border, as in the barrier-layer rectifiers, on a region of rapidly increasing intrinsic conductivity whose resistance in relation to the barrier-layer region can be neglected. On the contrary, the further layers possess no greater (neutral) defect density than the surface layer and there is consequently added to the surface-layer resistance a "path (series) resistance" which is proportional to the entire residual layer thickness of the semiconductor. A significant dependence on potential (and direction) of the total resistance is possible only when the surface resistance is large as against the path resistance, at least in the inverse direction. As theoretically not very different from this must be regarded the requirement that the zero resistance must already be appreciably greater than the path resistance.

If we restrict ourselves initially to plane arrangements, we can then define an equivalent layer thickness l^{o}_{aeq} which, possessing the intrinsic conductivity $\varkappa_H$ of the defect semiconductor, would have at zero bias precisely the same resistance as the surface layer *. If d is the total thickness

* In the more detailed calculations of the later publications, surface-layer resistance does not characterize the total resitance at a layer spatially limited somehow but the difference -- independent of the selection of this boundary -- between total resistance and the resistance of a semiconductor layer with the conductivity $\varkappa_{II}$ assumed as extending up to the metal boundary. If the surface-layer resistance .../...

of the semiconductor, then the requirement of precedence of the surface resistance over the path resistance requires that the following condition must be satisfied:

$$d \ll l^{o}_{aeq}.$$

The equivalent thickness l^{o}_{aeq} can be calculated from the surface-layer theory. It is shown to be dependent on the surface density n_I at the respective electrode as well as on the dielectric constant of the semiconductor and on temperature** but is independent of the intrinsic conductivity of the semiconductor. Since the surface layer is a zone of impoverishment which possesses higher specific resistance in order of magnitude than the massive semiconductor, the equivalent thickness is moreover appreciably greater than, for example, the zone of significant surface impoverishment marked by n less than $1/2_{n_H}$. For n_I values corresponding to a surface conductivity on the order of magnitude 10^{-9} Ohm^{-1} cm^{-1} l^{o}_{aeq} results in order of magnitude as 10^{-2} cm and for surface conductivities of about 10^{-5} Ohm^{-11} cm^{-1} as 10^{-4} cm.

---

* .../... is large against that of the semiconductor layer of equal thickness, then there exists no appreciable difference between "total"and"additional" barrier resistance.

**There exists a further dependence on the number of possibly existing electron traps.

The first case appears to correspond approximately to the conditions in cuprous oxide (of arbitrary conductivity), in contact with copper, gold, etc. We would therefore, e.g. in cuprous-oxide plates about 1/10 mm thick whose defect volume is homogeneous up to the metal boundary (diffusion of the electrodes at very low temperature), not yet need to expect any significant predomination of a non-linear surface resistance over the path resistance. This corresponds to the observations on massive cuprous-oxide plates, carried out at room temperature in contact with various metallic electrodes diffused at low temperature. Also in agreement with the theory is the observation that, at low-temperature measurements where we must assume significantly smaller n_I-values and correspondingly greater 1^{o}_{aeq}-values, a resistance depending on potential is observed also for such "non-blocking"electrodes on 0.1-mm thick cuprous-oxide plates of arbitrary intrinsic conductivity.*.

The occurrence of such non-linear effects in homogeneous semiconductor layers of very minor thickness with "inversion-free" applied contacts has so far not found any

---

* F.Waibel,1933,unpublished.The transfer resistances at electrodes observed in cuprous-oxide strips at -180° by Nasledow and Nemenow, Sow.Phys.7,513,1935 do not appear, however, to belong here directly since, at the high fields utilized, polarization effects through ion movement at the electrpdes are not excluded.

special interest and therefore also no systematic experimental treatment. There is no doubt that much remains to be done experimentally in this field. For a long time and under certain viewpoints even today, interest was generally high in the rectifier action of semiconductor-metal contacts where we can express this action as the reduction through a diffusion process to small amounts of the effective path length, i.e. the "detector" arrangement of metallic point and semiconductor.

If r_K is the radius of the contact surface between metallic point and semiconductor assumed as circular ("contact radius") then, under homogeneous conductivity $\varkappa$, we know that the diffusion resistance is equal to $1/\varkappa \cdot 1/4 r_K$. Since the surface of contact is πr_K^2, an effective path length of $\frac{\pi r_K^2}{4 r_K} = \frac{\pi}{4} r_K$ i.e. of approximately equal magnitude as the contact radius, corresponds to the diffusion resistance.

Instead of the above relation between l^o_{aeq} and d, we are therefore now able to postulate, as condition for the predomination of the surface resistance over the path resistance, a relation

$$r_K \ll l^o_{aeq}$$

where l^o_{aeq} is to be calculated as before for the plane arrangement.

This postulate, confirmed by more detailed considerations,

actually agrees most astonishingly with the actual observations, on point-contact rectifiers where the dependence of rectifier action on magnitude of contact surface was investigated. From the publications of Holm* we are quite adequately familiar with the contact surfaces occurring when bringing metal points in contact with (hard) crystals. With sufficient pointed metal contacts, circumstances here are so that the contact surfaces formed through the elastic yielding of the metal are too small in order to support the applied pressure without exceeding the yield point. Even at very low contact pressures, the metal of the point begins to yield at the contact surface and the latter is enlarged until the specific yield pressure is attained or falls short. Under these assumptions, there results contact surfaces with diameters of about 10^{-3} cm under the normal pressure of about 10 g for point contacts. Under higher contact pressures, the contact surface changes (irreversibly) proportional to the contact pressure i.e. the contact diameter changes proportional to the root of the contact pressure. If thesurface of the semiconductor, e.g. through destruction of an impurity layer, were to be damaged by the applied metal point, then this would have to be, within the yield-pressure range, exactly the same case with low as with high contact pressures because the specific load on the surface here does not change. A change of the surface state of the semiconductor in response to contact pressure accordingly is rather improbable. In spite of this, we find generally a failure

* General interrelations,Wiss.Veroeffentl.Siemens Konzern 7,H2,217 1929;detailed data e.g.from R.Holm u.W.Meissner,ZS f.Phys.74, S.272,1932

of the point-contact detectors at higher contact pressures and this failure occurs in order of magnitude at pressures which would correspond, according to Holm, to a contact surface with a diameter of about 10^{-2} cm. These observations can be interpreted obviously without difficulty if we assume that these are precisely the diameters at which the equivalent thickness l^{o}_{aeq} of the respective surface-layer resistance reaches the same order of magnitude as the contact radius r_K controlling the ohmic diffusion resistance. The order of magnitude of l^{o}_{aeq} to be assumed from this actually does agree in order of magnitude with the above theoretical data on l^{o}_{aeq}.

Systematic experimental investigations will produce further information here also. Specifically, it should be possible to show that, with decreasing temperature (i.e. decreasing surface density n_I), the permissible contact pressures may increase whereas, at higher temperatures, only very loose contact will still permit manifestation of significant detector effects. However, even without new and more exact experimentation in this direction, we need not hesitate to enter the described concordance in order of magnitude between the calculations of the surface-layer theory and the observations on point-contact detectors as a credit to the space-charge theory.

Let us stress at this point once more that it is not the

existence of a given equivalent length which should be regarded as a specific conclusion from the surface-layer theory. In any theory and any experiment which demonstrates an anomalous transfer of resistance at the locus of contact between point and semiconductor, a substitute path in the semiconductor itself can be coordinated with this transfer resistance which produces the same resistance as the point of transit. Characteristic for the surface-layer theory are on the contrary the specific indications on equivalent length and specifically the indication that in the absence of special barrier layers, only equivalent lengths of microscopic dimension can be expected which automatically lead to a subordination of the unipolar resistance to the ohmic diffusion resistance as soon as the point-contact surface exceeds given dimensions. If we proceed from the general statement that, for the formation of special chemical barrier layers (defect impoverishment) at the boundary between metal and semiconductor, extraordinary prerequisites are necessary and that a defect density relatively homogeneous up to the surface is the normal case, then we obviously understand, within the frame of the surface-layer theory also the previously somewhat mysterious fact that there are very many kinds of point-contact detectors but only very few "real" rectifiers. This is precisely due to the fact that it is possible without difficulty to realise, in point-contact arrangements where the contact radius instead of the layer thickness

is controlling for the path resistance, only those microscopic dimensions of the path resistance connected in series with the surface resistance which permit with homogeneous defect density a predominance of the surface resistance.

8. Other Semiconductor Arrangements with Potential-Dependent Resistance

In addition to the customary detector arrangement with metal-point contact, there are without doubt a number of other arrangements in which the ratio of path resistance can be reduced so that a high potential-dependence of the total resistance, frequently in connection with high rectification effect, is achieved. Together with Stoermer and Waibel (ref. 9), I devoted to this question of "geometric rectification" a special investigation in 1931 whose conclusions in regard to the universal significance of such effects subsequently could not be maintained even for grown or diffused electrodes. However, for all impressed electrodes (copper and graphite points, impressed foil, coating with hydrogen layers of low pressure head, influence of roughing of the protoxide surface,etc.), the discussion in this paper which was also based on the joint action of a surface effect with an interposed path resistance, should have on the whole brought out the essentials. The observed increase of the relative rectifier action when we compared ruled gold electrodes (atomized on

cuprous oxide) with a width of 10^{-2} cm with those of a width of 10^{-4} cm, can now be utilized even quantitatively in favor of the surface-layer theory.

Within the same framework belongs in my opinion -- so that we should enter here on this important question -- the observed rectifier action on semiconductors which are separated from the metal electrodes in contact with them by a so-called "artificial barrier layer". Diffusion of quartz layers with a thickness of 10^{-6} to 10^{-5} cm on massive cuprous-oxide plates originally permitted observations of strong blocking and rectifier action (ref.10) which could possibly be still explained by transit of electrons through the artificial layer. These artificial-layer effects became more and more mysterious after it was shown that any arbitrary insulating layers -- even organic lacquer coatings -- with a thickness up to 5 μ and beyond (ref.11) displayed similar effects, although with higher absolute resistances. A satisfactory explanation of this effect was probably given by the investigations of Klarmann and Mühlenpfordt (ref.12) who demonstrated that there exists, even in artificial barrier layers produced by the most perfect methods, microscopic fissures or holes -- in lacquer films, pore-like holes with a diameter of 1 μ -- through which the subsequently diffused metal electrode comes into "point-like" direct contact with the semiconductor. We thus do not commit an error if we regard all "artificial-

layer rectifiers", or at least those with artificla layers more than 10^{-6} cm thick as (multi-)point-contact rectifiers with very thin point contacts spaced far apart. That under these circumstances the path resistance in relation to the surface-layer resistance is particularly low and that we can therefore expect particularly efficient rectification agrees with observations. In particular, for some excess semiconductors in which it was previously impossible to attain any pronounced rectifier of very thin layer thickness, this method led to efficiency ratios of 1:100 to 1:1,000 at potentials of +/-1 Volt (ref.13).

With the discussion of the point-contact rectifiers based on the assumption of homogeneous defect density in the semiconductor in this and the preceding chapter, it was not intended to state that, in semiconductor surfaces with point contats, defect in homogeneity, i.e. chemical barrier layers in the mere sense of section 5, may not also play a role. We merely desire to point out that they do not necessarily play a role. There will probably be also cases where (invisible) high-ohmic impurity layers exist either on the semiconductor surface or on the metal electrode and act similar to artificial layers, i.e. have no other purpose than to make the real surface of contact significantly smaller than would correspond to the normal relations between contact pressure and contact surface metnioned above.

Finally, we should probably point out already here that

an influence on the thermal surface density n_I -- assumed so far as defect independent -- and consequently of the equivalent length l^o_{aeq} by the defect volume, perhaps changing from locus to locus, on the surface of the semiconductor must be taken into account specifically then when , e.g. as in galena, the defect volume is so high that conductivity above 1 Ohm^{-1} cm^{-1} occurs on the surface. Under these circumstances, the surface-layer theory lets us expect, even without bias, the occurrence of very high mean and even higher local surface fields. So-to-speak, the metal is no longer leakproof and electron impoverishment on the surface ceases. We will therefore not be surprised to find no longer any point-contact rectification in freely conducting galena specimens with conductivities on the order of 10^3 Ohm^{-1} cm^{-1}.

The space-charge theory explains in very easily comprehensible manner the rectifier effects occurring at the point of contact of semiconductor specimens of the same kind but with different defect volume, pressed against each other. These effects were previously difficult to understand. They occur, for example, under contact of evacuated with oxygen-rich cuprous oxide but even under contact of cuprous-oxide specimens polished or sanded in air (and thus superficially slightly enriched with oxygen) and acid-etched, i.e. superficially somewhat impoverished in oxygen (ref. 14). Within the framework

of the space-charge theory, we will assume that conditions here are similar to those at the n_1^x-n_2^x boundary within a semiconductor as in fig. 9. Forward direction is accordingly the direction of the (defect-)electron transfer from the freely to the poorly conducting cuprous oxide* in agreement with observation. That fresh fracture faces of thin cuprous-oxide planes in point-contact show only a pure ohmic diffusion resistance (Stoermer 1930) is of interest for the surface-layer theory since we can conclude, at least in this one case, reliably on a defect volume on such surface homogeneous up to the surface. Any disturbance extending to finite depth would have as consequence, if not unipolarity, a non-linearity of the observed resistance.

The dependence on potential of the resistance of massive and relatively freely conducting crystal fragments (specifically

---

* Where only superficial changes of the defect volume are concerned, there must be taken into account, in addition to the defect jump, also the defect changes in transition from anomalous to normal defect density in the interior of the two specimens; however, since these transitions are slower than at the contact point of the specimens, they have a more quasi-neutral character and consequently condition lesser rectifier effects. Considerations pointing to a supposed interrelation between the steepness of the changes of property of a layer and the magnitude and direction of its rectifier effects are found already in F. Waibel,Wiss. Veroeff. Siemens-Werke 15, H.3.S.75, 1936.

selenium, boron, silicium carbide * would contradict the space-charge theory if in these cases either the homogeneity or an only gradual spatial variability of the defect volume of the respective layers could be demonstrated. However, all more recent observations point in the opposite direction. The non-linear processes appear to be restricted to narrow layers in the vicinity of internal surfaces or impurity zones and thus evidently take place in "space-charge dimensions" as we have shown them to be probable for boundary-surface rectification. In this field also, experimentation will have to furnish further information, e.g. through probe scanning of the potential gradient in such energized crystals for the demonstration of discrete high-ohmic layers.

Under particular conditions, however, the layers of anomalous conductivity increase under the space-charge theory also to macroscopic dimensions. As characteristic length of a surface layer was given in Section 4 a parameter l_o (surface-layer dimensions which should not be confused with the generally significantly larger equivalent length l^o_{aeq}) which attains at an intrinsic conductivity $\varkappa_H$ of the semiconductor of 10^{-6} Ohm^{-1} cm^{-1}, the amount of about 10^{-3} cm = 10 μ. If we now investigate very high-ohmic semiconductors, e.g. $\varkappa_H$ about 10^{-10}

---

* (See ref. 15)

Ohm^{-1} cm^{-1}, then l_o becomes larger at a ratio of $1/\sqrt{\varkappa_{II}}$ and the surface layer -- and any other disturbance of the quasi-neutral state -- will equalize itself therefore only within a distance of about 10^{-1} cm = 1 mm and, within lesser layer thickness, there can be no expectation of a quasi-neutral state and ohmic resistance.* The investigation of such cases is also a task for the future. It is somewhat hindered experimentally by the high resistance of the arrangement and theoretically through the no longer additive interaction of the two electrode effects. In general, a quasi-neutral zone then does not occur within the semiconductor.

We arrive at simple and clear conditions again only when the layer thickness becomes so small in relation to the surface-layer thickness l_o that we can assume constant field intensity within the entire layer. This corresponds to the case of

---

* The dimensions here given are valid, however, only on the assumption that the semiconductor does not contain any electron traps, i.e. that the electrons are capable of existing either only free or in conjunction with their positive defect partners but not as (transiently) immobile negative charges in the lattice (correspondingly applicable to defect electrons). As will be shown later, otherwise all dimensions are shortened at a ratio of $1/\sqrt{f}$ where f signifies the ratio of the entire lifespan of the electron to their lifespan in the free state. As far as can be shown indirectly, no traps appear to occur in normal cuprous oxide. On the other hand, I believe it possible that, in the cuprous oxide evacuated very highly and for a very prolonged time as used by F.Waibel for his investigations on microscopic barrier layers (also elsewhere, ref. p 63), there were created oxygen gaps in addition to copper gaps which have positive fixed charges decisively influencing the space-charge distribution.

fig. 2, "semiconductor as vacuum", where the distribution of electron density is controlled only by the surface density determined by the metal electrodes and the externally applied field (together with the intrinsic field of the diffusion potential). In contrast to the "thin-film rectifiers" which satisfy only the condition $d \ll l_{Leq}^{0}$, these arrangements satisfying the condition $d \ll l_0$ will be most suitably designated as "potential-barrier arrangements" and, in as far as they are limited by electrodes with different work function, as "potential-barrier rectifiers". We shall return in future publications to the special significance of such arrangements for investigation of the undisturbed thermal work function metal-semiconductor.

This is not the place to discuss other experimental confirmations of the space-charge and surface-layer theory of rectification which can be done only in a later communication subsequent to the conclusion of the required calculations. The most important accomplishments of the theory in addition to the date on order of magnitude described in Section 7 and 8, result already from Section 3 to 5, i.e. the unconstrained explanation of the steady transition from high surface resistances in the inverse direction to vanishing surface resistances in the forward direction as well as the related changes of capacity as a consequence of the steadily increasing contraction of an impo-

verished surface layer on the effective electrode and an equally unconstrained interpretation of the actually observed sense of direction of rectifier action as a consequence of the character of excess or defect conduction of the interior of the semiconductor.

9. Historical Remarks

The inducement for abandoning the gap theory of detector action was given by the discovery and investigation of surface rectifiers, in particular that of the cuprous-oxide rectifier developed by L.O. Grondahl * between 1920 and 1924 and repeatedly described and discussed since 1925. Grondahl already suspected that the interaction of field and diffusion effects in a layer of finite thickness within the protoxide controlled the unipolar processes.

Further theoretical attempts at interpretation in subsequent years were reported in various publications. I shall here merely refer to these communications** and restrict myself here to further report some facts on the origin of the presently existing theory which was here evidently described only qualitatively. I shall have to stress in particular the contributions of R. Peierls to the development of the new rectifier theory in correspondence with me from 1931 to 1933.

* USA-Pat.1640335 of 7 Jan.1925;Sience 36,306,1926;Summary,Rev Mod.Phys.5,141,1933

** e.g. N.F.Mott

If I wrote, during the discussion of the capacity measurements on cuprous-oxide rectifiers in 1929 (ref. 16) that "we may perhaps imagine that the presence of excess electrons in the boundary layer greatly reduces conductivity because so-to-speak the positions available for the conduction electrons are occupied by statistical electron space charges and thus become clogged for the transit of the conduction electrons", then this is already rather close to the present space-charge theory of the surface layer if we consider that cuprous oxide is a defect semiconductor and that here actually, in the zone of surface impoverishment, the presence of negative space charges works hand in hand with an impoverishment of defect electrons, i.e. a clogging of the electron gaps of the highest still filled energy band. Actually, this remark (which was made prior to the wave-mechanic concept of the whole conduction mechanism in semiconductors;) explained only the action of a space charge (assumed as existing) on the surface conductivity but not really the origin of this space charge. Although I find, in my papers of Dec. 1931, differential equations for the fields and particle numbers which were developed in consideration of the field and diffusion currents, the space-charge conditions and defect equilibria and are entirely identical with the present equations of the surface-layer theory, there was still absent clear recognition of the fact that surface impoverishment conditioned through the vicinity of the metal

in the semiconductor at the metal boundary must be regarded as the primary cause of the entire space-charge surface layer.

It therefore constituted without doubt decisive progress when Peierls who had arrived at similar differential equations in our joint effort for clarification of theproblem, wrote in a letter of March 4, 1932 as follows: ... "The essential factor is the condition for n" (our surface density n_R). The latter does not necessarily need to be such that the density of the semiconductor becomes equal to the electron density in the metal; this "surface density" actually does not need to be greater than n_o" (our n_H) but can also be less. It actually must be less if the contact potential between metal and semiconductor has a suitable sign. It then appears to be a very obvious and effortless assumption that, in the practically existing detectors, this surface density becomes less than n_o so that a layer of lower electron density, i.e. high resistance, occurs on the surface".

Instead of the "contact potential" which is utilized generally only as a measure of the purely electrostatic potential differences at the boundary between two conductors, we probably have to regard today the "work function" of the electrons, dependent on a chemical and electric component, as the parameter

controlling n_r*. Not the sign but the amount of this parameter appears to be controlling; for the difference between surface density n_R and internal density n_H, the defect character of the semiconductor plays a role not then recognized**.

Nevertheless, in this quotation from Peierls, the basic concept of a surface impoverishment impressed on the semiconductor by the metal is clearly expressed. A second and extremely important point is that the concept of surface impoverishment due to the adjacent metal can now be expanded without diffi-

---

* Cf. Remarks on work function in Section 2. Decisive for the rectifier theory of semiconductors is the additional determination that this work function is independent of the defect volume and consequently of the characteristic conductivity of the semiconductor and has no relation to the electric potential difference created between the metal and the interior of the defect semiconductor. Cf. R. Fowler, Statistical Mechanics, Cambridge 1936, and the contradictory discussion by N.F. Mott, elsewhere, note on p. 27, 1939. The author utilized the concept of defect-independent work function time in the investigation of electron transition through crystal layers without characteristic conductivity for the first time in 1934.

** Note during manuscript correction: In answer to a question, I received in the meantime a response from Peierls that he understood at that time, by "contact potential", the potential difference established from the Boltzmann principle by the ratio between surface density and internal density. This potential between the interior and the boundary of the semiconductor occurring in thermal equilibrium ("diffusion potential") actually changes sign with $n^R \gtrless n_H$ but must be regarded rather as an associated manifestation of surface impoverishment than as a prediction of the conditions of its occurrence.

culty to excess semiconductors and thus ads to the supposition of the interrelation between conduction character and direction of rectification which was so well confirmed by the investigations of pag.28 on excess semiconductors.

Accordingly, we had come very close to our present understanding already in the spring of 1932. However, various strenuous detours were required before it was possible to further specify the earlier basic concepts and to apply them in confidence to the rectifier processes.

Two objections confronted the interpretation desired by us. The first was the statement of the sign of the Hall effect in cuprous oxide which permitted conclusions of excess (electron) conduction at room temperature and would thus have produced, within the framework of our theory, the wrong sign for the direction of rectification. The other was the determination that significant transfer resistances and rectifier action appeared to occur only under special pretreatment (i.e. apparently extensive defect impoverishment) of the cuprous-oxide layer adjacent to the metal. These two objections could not be countered at the time and prevented any publication on our part (1949).

Consequently, the old potential-barrier theory in various modifications dominated the field then and in subsequent years. This theory was able to convincingly substantiate both

the orientation of the rectifier action in cuprous oxide (however, with the erroneous assumption of excess conduction;) and the necessity of a special state of the protoxide surface. This applied both to the theory suggested by van Geel (ref. 17) according to which the electrons can penetrate the barrier not due their thermal energy but only under the action of strong fields as well as for the theories published approximately simultaneously, of Frenkel and Joffé (ref. 18), Wilson (ref. 19), and Nordheim (ref. 20) in which the potential barrier is penetrated without applied field wave-mechanically by electrons of both directions and the applied field changes only the levels on either side of the potential barrier and thus influences the currents traversing in both directions.

That high space-charge densities within the semiconductor are necessary for creating such changes of level of a potential barrier as well as for formation of the latter in the no-current state which cannot immediately be made compatible with the properties of the semiconductor, was not overlooked in particular by Nordheim who identified the width of the potential barrier with the width of an oxygen-impoverished layer in cuprous oxide. In the discussion of this problem by correspondence, however, Nordheim was then still able to point quite correctly to the character of excess conduction of cuprous oxide which would have produced the wrong sense of orientation within the framework of our theory.

After Fritsch in 1934 corrected the sign of the Hall effect in cuprous oxide and thus demonstrated the character of hole conduction of this standard substance, this objection might have been considered as refuted and the superiority of the space-charge theory as confirmed. However, the significance of the defect density of the semiconductor for the absolute magnitude of the surface-layer resistances (fig. 4, Section 3 herein) had not been clearly recognized so that the second objection concerning the entirely different behavior of the same metal electrodes in contact with semiconductor layers of different defect volume was retained unchanged.

It consequently proved necessary to await further experimental determinations and to admit also other theoretical possibilities. In this sense must be evaluated the theories pursued beginning in 1931, according to which the "blocking" and "non-blocking" contacts of the same metal with a given semiconductor were said to be differentiated by the fine structure of the contact* which could lead to differences in order of magnitude of the true contact surfaces. After this hypothesis whose

---

* (See ref. 21)

significance for impressed metal foil (artificial-layer rectifiers, etc.) was given due consideration in Section 8, was refuted through thermal-conduction experiments for the processes at the real wall of the protoxide layer grown on copper, the expectations placed in the space-charge theory necessarily reached a low level.

In 1935, more detailed investigations on rectifier action on wet and dry layers of aluminum-oxide rectifiers as well as some experiments on the direction of rectification in zinc oxide encouraged a communication * on the deduction of the sense of orientation of the space-charge theory, i.e. preferred transition of the conducting particles from the semiconductor into the electrode and consequently inverse sense of orientation of the rectifier action in excess and defect semiconductors. The further experiments with zinc oxide and titanium dioxide of Sec. 6 confirmed this supposition and thus restored a certain confidence in the space-charge theory.

However, another delay set in. Experiments of Waibel on superevacuated protoxide layers on copper ** seemed to indicate that the deviations from ohmic behavior of a semiconductor were after all more a question of field intensity and thus of the

---

* (see ref. 22)

** Elsewhere, ref.42.

mean velocity of the electrons than a consequence of their density capable of being influenced through space charges. These experiments certainly contain important indications and may perhaps probe fertile in another direction. Nowever, for normal rectifier defects, the supposed increase of mean free path* through field effects does seem to appear to play a very subordinate role and the concept of a quasi-thermal behavior is here most likely far more appropriate.

Only further experimental investigations of an apparently rather remote kind finally led back to the space-charge theory. Klarmann and Muelenpfordt** investigated current flow

---

* A.V. and A.F. Joffé justifiably objected to the assumption of an enlargement. It is surprising that, under these circumstances, the authors still appear to regard conductivity dependent on field strength, event hough in conjunction with space-charge effects, for the actual cause of rectifier action. If the mean free path, and consequently the mobility is not influenced by the field, there can then be concerned only the density of the electrons. However, the latter is influenced decisively not by certain universal and critical field strengths but by additional field strengths which are comparable with the not at all universal diffusion-field strengths within the respective surface layer.

** (see ref. 23)

through layers of quartz 10^{-5} cm thick between silver electrodes and here found a non-linear characteristic which appeared however, to transform into an ohmic resistance at the lowest potential. This last finding was checked in particular by Thomer* and definitely confirmed. Moreover, a significant temperature trend of this zero resistance was found. This induced the introduction of the concept of a purely thermally determined work function metal-semiconductor (and/or metal insulator) with greater confidence also into the semiconductor theory of the surface and point-contact rectifiers and developed the theory in he manner outlined above now sufficiently far so that it encompasses also the influence of the defect volume of the semiconductor on the magnitude of the transfer resistances.

In a further communication "On Quantitative Development of the Space-Charge and Surface-Layer Theory of Crystal Semiconductors", the mathematical formulation of the theory here developed will be given, initially on the assumption of homogeneous defect volume in the semiconductor and by disrregarding field effects. Although it is not possible, by reason of the interrelations occurring to represent the relation between current and voltage generally by known functions, the most important properties of the current-voltage curve can be derived analytically, specifically zero resistance, i.e. the guiding

---

* Unpublished investigation from the Research Laboratory of the Siemens Company.

constant in zero bias essential for semiconductor action, as well as the boundary laws in the forward and inverse direction.

This will disclose the identical character of the diffusion voltage existing in equilibrium with the "sluice voltage" necessary for the elimination of the barrier resistance in the forward direction.

There is also planned a further study on the question of electron transfer between metal and semiconductor in which we shall discuss the theoretical indications on the magnitude of this work function (similar to the thermionic-electron problem) for excess and defect electrons as well as point out the particular interrelations resulting from the existence of an internal equilibrium between excess and defect electrons in the semiconductor.

The influence on thermal work function through field effects requires further investigation and here also the findings of the vacuum theory cannot be simply transposed.

Notable data result from the calculation of capacity of a space-charge surface layer or, more generally, of the complex impedance of such a layer as a function of bias and frequency. The author wrote on this in collaboration with Spenke.

Undoubtedly, the clarification of the circumstances with weak fields will contribute also to a more exact evaluation of various interesting groups of manifestation under which a quasi-thermal behavior can no longer be assumed in the semiconductor. Among these problems are the electric breakdown of thin oxide layers, the "Malter" effects at the boundary between metal and semiconductor.

LITERATURE REFERENCES

1. Die Naturwissensch.26, 843, 1938.
2. F. Braun, Pogg. Ann. 153, 556, 1874.
3. W. Schottky, ZS. f. Phys. 14, 63, 1923
4. W. Schottky u. W. Deutschmann, Phys.ZS.30,839,1929.
5. O. Fritsch, Ann. d. Phys.22, 375, 1935.
6. A.H. Wilson, Proc. Roy, Soc. London (A)133, 458 and 134, 277,1931; R. Peirls, Erg.d.exakt.Naturwissensch. 11,264,1932.
7. F. Waibel, ZS. f. techn. Phys. 11,366,1935.
8. C. Wagner, ZS. phys.Chem.(B)21,25,1933.
9. Jahrb.d.drahtl.Telegr. 37,162,u. 175, 1931.
10. W. Jusé, Nature 132,242,1933.
11. W. Jusé, Sow. Phys.7.1,1935; W.Ch.van Geel u.J.H.de Boer, Physica 2, 892, 1935.
12. ZS. f. Elektrochem. 44, 603, 1938.
13. W. Hartmann, ZS.f.techn.Phys. 17, 436, 1936 fuer ZnO; H. Klarmann, Wiss. Veroeffentl. Siemens-Werke 18, H.2. S.198, 1939 fuer TiO_2.
14. R. Stoermer, 1930, unpublished.
15. F.P. Henninger, Ann. d. Phys. (5)28,245, 1937 u.31 337, 1938; cf.also for earlier literature.
16. W. Schottky u. W. Deutschman
17. ZS. f. Phys. 69, 765, 1931; ibid 87,220,1933 ; Physica 1,531, 1143, 1934; . Der Ingenieur 50,E9, 1935.

18. Sow. Phys. 1, 60, 1932.

19. Proc. Roy. Soc. London (A) 136, 487, 1932.

20. ZS. f. Phys. 75, 434, 1932.

21. W. Schottky, R. Stoermer u.F. Waibel, On rectification action at the boundary between cuprous oxide and applied metal electrodes. Jahrb. f. drahtl. Telegr. und Teleph. 37,162 u. 175, 1931.

22. ZS.f. Phys. 75,434,1932.

23. On Electron Transition through Thin Insulating Layers, ZS.f. Elektrochem. 44,603,1938.